SYSTEMS AND METHODS FOR WORD PREDICTION

AND SPEECH RECOGNITION

Claim for Priority

This application claims priority from Japanese Patent Application No. 2000-
5   210599, filed on July 11, 2000, and which is hereby incorporated by reference as if fully
set forth herein.

Field of the Invention

The present invention relates to a speech recognition apparatus and a method
therefor, and in particular to a speech recognition apparatus for recognizing the natural
10   language spoken by persons that thereafter is used for composing sentences and for
creating text data and a method therefor.

Background of the Invention

A statistical method for using an acoustic model and a language model for
speech recognition is well known, and has been featured in such publications as: "A
15   Maximum Likelihood Approach to Continuous Speech Recognition," L.R. Bahl, et. al.,
IEEE Trans. Vol. PAMI-5, No. 2, March, 1983; and "Word based approach to large-

vocabulary continuous speech recognition for Japanese," Nishimura, et. al., Information

Processing Institute Thesis, Vol. 40, No. 4, April, 1999.

According to an overview of this method, a word sequence W is voiced as a

generated sentence and is processed by an acoustic processor, and from a signal that is

5   produced a feature value X is extracted. Then, using the feature value X and the word

sequence W, assumed optimal recognition results W' are output in accordance with the

following equation to form a sentence. That is, a word sequence such that, when the

word sequence W is voiced, the product of the appearance probability P (XW), of the

feature value (X), and the appearance probability (P(W)), of the word sequence W, is

10   the maximum (argmax) and is selected as the recognition results W'.

[Equation 1]

$$W' = \overset{\text{arg max}}{\underset{w}{}} P(W \mid X) \overset{\text{arg max}}{\underset{w}{}} P(w)P(X \mid W)$$

where P(W) is for a language model, and P(X W) is for an acoustic model.

In this equation, the acoustic model is employed to obtain the probability

15   P(X|W), and words having a high probability are selected as a proposed word for

recognition. This language model is frequently used to provide an approximation of the probability P(W).

For the conventional language model, normally, the closest word sequence is used as a history. An example is an N-gram model. With this method, an approximation of a complete sentence is produced by using the probability of the appearance of N sequential words, i.e., an approximation of the appearance probability of the word sequence W. This method is exemplified by the following established form.

[Equation 2]

$$P(w) = P(w_0)P(w_1 \mid w_0)P(w_2 \mid w_0w_1), ..., P(w_n \mid w_0w_1, ..., w_{n-1})$$
$$= P(w_0)P(w_1 \mid w_0) \prod_{i=2}^{n} P(w_i \mid w_{i-2}w_{i-1})$$

Assume that in the above equation the appearance probability of the next word W[n] is affected only by the immediately preceding N-1 words. For this purpose, various values can be used for N, but since N=3 is frequently employed because of the balance it provides between effectiveness and the learning data that is required, in this equation, N=3 is employed, and the above method is therefore called a tri-gram or a 3-gram method. Hereinafter, when the n-th word in a word sequence W consisting of n

words is represented by W[n], the appearance probability condition for the calculation

of the word W[n] is that there are N-1 preceding words (two words), i.e., the

appearance probability for the word sequence W is calculated using P(W[n] W[n-

2]W[n-1]). In this equation, the statement to the left (W[n]) of " " represents a word to

5   be predicted (or recognized), and the statement to the right (W[n-2]W[n-1]) represents

the first and the second preceding words required to establish the condition. This

appearance probability P(W[n] W[n-2]W[n- 1]) is learned for each word W[n] by

using text data that have previously been prepared and stored as part of a dictionary

database. For example, for the probability that a "word" will appear at the beginning of

10   a sentence, 0.0021 is stored, and for the probability a "search" will follow, 0.001 is

stored.

The Tri-gram model will now be described by using a simple phrase. This

phrase is "sara-ni sho-senkyoku no (further, small electoral districts)" and is used to

predict the following "donyu (are introduced)". Fig. 8A is a diagram showing the state

15   before the prediction is fulfilled, and Fig. 8B is a diagram showing the state after the

prediction is fulfilled. As is shown in Fig. 8A, the phrase consists of five words, "sara-

ni", "sho", "senkyo", "ku" and "no", while the predicted word is represented by "?", and

the arrows in Figs. 8A and 8B are used to delineate the modifications applied to the

words. As previously described, in the tri-gram model, two preceding words are constantly employed to predict a following word. Therefore, in this example, "donyu" is predicted by "ku" and "no", words enclosed by solid lines in Fig. 8A.

However, depending on the sentence structure, the tri-gram method for
5   employing two immediate words to predict a following word is not the most appropriate. For example, the tri-gram method is not appropriate for the case illustrated in Fig. 9, wherein the phrase "nani-ga ima seiji-no saisei-no tame-ni (at present, for reconstruction of the politics, what)" is used to predict a word. According to the tri-gram method, as is shown in Fig. 9A, "tame" and "ni" are employed to predict
10   "hitsuyo (is required)". But in addition to these words, other structurally related words, such as "nani" or "ima" must be taken into account in order to increase the accuracy of the prediction.

Chelba and Jelinek proposed a model for employing the head word of two immediately preceding partial analysis trees to predict a succeeding word. According
15   to the Chelba & Jelinek model, the words are predicted in order, as they appear. Therefore, when the i-th word is to be predicted, the (i-1)th word and the structure are established. In this state, first, the head word of the two immediately preceding partial analysis trees are employed to predict, in the named order, the following word and its

JP920000133US1   - 5 -

speech part. At this time, the modification relationship between the head word of the

two immediately preceding partial analysis trees and the predicted word is not taken

into account. After the word is predicted, the sentence structure that includes the word

is updated. Therefore, the accuracy of the prediction can be improved compared with

5   the tri-gram method, which employs two immediately preceding words to predict a

following word. However, in the model proposed by Chelba and Jelinek, a word is

predicted by referring to the head word of the two immediately preceding partial

analysis trees, regardless of how the words are modified, so that, depending on the

sentence structure, the accuracy of the prediction may be reduced. This will be

10   explained by referring to the phrase "sara-ni sho-senkyoku no", used for the tri-gram

model.

As is shown in Figs. 10A to 10C, the phrase "sara-ni sho- senkyoku no" is

constituted by two partial analysis trees, and the head word of the trees are "sara-ni"

and "no", which are enclosed by solid lines in Fig. 10A. Therefore, according to the

15   method proposed by Chelba and Jelinek, "sara-ni" and "no", which are two

immediately preceding head word as is shown in Fig. 10B, are employed to predict the

next word "donyu". When "donyu" is predicted, as is shown in Fig. 10C, the sentence

structure including "donyu" is predicted. In the prediction of the structure, the

modification of words as indicted by arrows is included. Since "sara-ni" does not

modify "donyu", it is not only useless for the prediction of the word "donyu", but also

may tend to degrade the prediction accuracy.

For the phrase "nani-ga ima seiji-no saisei-no tame-ni", in Fig. 11, the following

5   prediction process is performed. This phase is constituted by three partial analysis

trees "nani-ga", "ima" and "seiji-no saisei-no tame-ni", and the head word of the trees

are "ga", "ima" and "ni". As indicated by the solid line enclosures in Fig. 11A, the two

immediately preceding head word are "ima" and "ni". Therefore, as is shown in Fig.

11B, "hitsuyo" is predicted by using "ima" and "ni". And after "hitsuyo" is predicted,

10   the sentence structure that includes "hitsuyo" is predicted, as is shown in Fig. 11C.

To predict a word, the modifications performed by words provides useful

information. However, that "nani-ga" is a modifier is not taken into account. As is

described above, according to the method proposed by Chelba and Jelinek, no

consideration is given for information that is useful for prediction that frequently

15   occurs.

A need therefore exists for a word prediction method and apparatus therefor that

supply improved prediction accuracy, and a speech recognition method and an

apparatus therefor.

Summary of the Invention

The present invention focuses on the fact that, at each word prediction step, a sequence of partial analysis trees covering currently obtained word sequences can be employed as historical information. A partial analysis tree sequence, when used as historical information, can be employed to select a partial analysis tree carrying information that can more usefully be employed for the prediction of the next word. In essence, when a word sequence employed as history and a modification structure are used to select the most useful word and/or word sequence for predicting the next word, prediction accuracy can be improved. That is, after a partial analysis tree that includes a modification function for a word to be predicted is specified, this partial analysis tree, i.e., a word and/or a word sequence that is estimated to have a modification relationship with a word that is to be predicted, is employed for the prediction of the following word. Unlike the method proposed by Chelba and Jelinek, since the structure of a sentence, to include the word to be predicted, is employed, only information that is useful for prediction will be taken into account.

Based on the above described idea, according to the present invention, a word prediction method is provided, said method comprising the steps of: specifying a sentence structure consisting of multiple words, including a target word to be predicted;

employing said sentence structure to select a word and/or a word sequence that has a modification relationship with said target word to be predicted; and predicting said target word based on said word and/or said word sequence that is selected. According to the present invention, it is preferable that the word and/or the word sequence

5   constitute a partial analysis tree in the sentence structure. According to the present invention, when multiple words and/or word sequences are selected, word prediction is performed based on the words and/or word sequences that are selected.

According to the present invention, a word prediction method is also provided, said method comprising the steps of: specifying a modification of a word to be

10   predicted by a different word and/or word sequence; and prediction the word using the different word and/or word sequence that is specified. According to the word prediction method, the modification includes a modification direction, and the word to be predicted may modify a prior word. When multiple modifications are established between the word to be predicted and the different word and/or word sequence, a word

15   can be predicted for each of the modifications.

According to the present invention, a speech recognition method is provided, said method comprising the steps of: specifying the structure of a phrase from the beginning of a sentence to the j-th word (j=0, 1, 2, ...); employing a sentence structure

up to the j-th word to specify one or multiple partial analysis trees modifying the

(j+1)th word; predicting the (j+1)th word based on the one or multiple partial analysis

trees; obtaining a sentence for a phrase including the predicted (j+1)th word and the

probability value for the sentence structure; and performing the above steps up to the

5   last word of the sentence, and selecting, as speech recognition results, a sentence

structure and a word sequence having maximum probability values.

Further, according to the present invention, a word prediction apparatus is

provided, said apparatus comprising: a storer which stores a dictionary wherein text

data that have been learned is written; and a predictor which employs a structure of a

10   sentence consisting of multiple words, including a word to be predicted, to select a

word and/or a word sequence that has a modification relationship with said word to be

predicted, and which predicts said word based on said word and/or said word sequence

that is selected.

According to the present invention, a word prediction apparatus is also

15   provided, said apparatus comprising: a storer which stores a dictionary wherein text

data that have been learned is written; and a predictor which selects a word and/or a

word sequence that has a modification relationship with a word to be predicted, and

which predicts said word based on said word and/or said word sequence that is

selected. According to the word prediction apparatus of the invention, for a predetermined word or word sequence based on the text data, a frequency whereat another predetermined word appears is stored in the dictionary. The predictor can calculate a probability value for a word sequence including a word that is predicted

5   based on the frequency. Further, the predictor can select, as a prediction result, a word sequence having the maximum probability value.

According to the present invention, a speech recognition apparatus is provided, said apparatus comprising: an acoustic processor which converts an input analog speech signal into a digital signal; a first storer which stores an acoustic model that has

10   learned a feature of speech; a second storer which stores a dictionary wherein an appearance frequency of a predetermined word relative to another predetermined word and/or word sequence is written; and a recognizer which uses said acoustic model and said dictionary to calculate a probability value for said digital signal, and which recognizes a word having the maximum probability value as input speech, wherein said

15   recognizer predicts a word to be predicted based on a structure of a sentence including said word, and employs said appearance frequency to calculate said probability value for said sentence, including said word that is predicted.

According to the present invention, a computer system is provided, said computer system comprising: a specifier which specifies a modification relationship between a word to be predicted and another word and/or word sequence, and which predicts said word by employing said word and/or word sequence modifying said word; and a display which displays said word that is predicted, and said modification relationship.

According to the present invention, a storage medium on which a computer readable program is stored is provided, the computer program permitting a computer to perform: a first process for specifying a sentence structure consisting of multiple words, including a target word to be predicted; a second process for employing said sentence structure to select a word and/or a word sequence that has a modification relationship with said target word to be predicted; and a third process for predicting said target word based on said word and/or said word sequence that is selected.

According to the present invention, a storage medium on which a computer readable program is stored is provided, the computer program permitting a computer to perform: a first process for specifying a modification relationship between a word to be predicted and a different word and/or word sequence; and a second process for predicting said word using said different word and/or word sequence that is specified.

According to the present invention, a program transmission apparatus is provided, said apparatus comprising: a storer which stores a program permitting a computer to perform a first process for specifying a sentence structure of a multiple word sentence, including a target word to be predicted, a second process for employing said sentence structure to select a word and/or a word sequence that has a modification relationship with said target word to be predicted, and a third process for predicting said target word based on said word and/or said word sequence that is selected; and a transmitter which reads said program from said storer and transmits said program.

According to the present invention, a program transmission apparatus is provided, said apparatus comprising: a storer which stores a program permitting a computer to perform a first process for specifying a modification relationship between a word to be predicted and a different word and/or word sequence, and a second process for predicting said word using said different word and/or word sequence that is specified; and a transmitter which reads said program from said storer and transmits said program.

According to the present invention, a program storage device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for word prediction, is provided, said method comprising the steps

of: specifying a sentence structure consisting of multiple words, including a target word to be predicted; employing said sentence structure to select a word and/or a word sequence that has a modification relationship with said target word to be predicted; and predicting said target word based on said word and/or said word sequence that is

5   selected.

According to the present invention, a program storage device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for word prediction is provided, said method comprising the steps of: specifying a modification of a word to be predicted by a different word and/or word

10   sequence; and prediction said word using said different word and/or word sequence that is specified.

According to the present invention, a program storage device readable by machine, tangibly embodying a program of instructions executable by the machine to perform method steps for speech recognition is provided, said method comprising the

15   steps of: specifying the structure of a phrase from the beginning of a sentence to the j-th word (j=0, 1, 2, ...); employing a sentence structure up to said j-th word to specify one or multiple partial analysis trees modifying the (j+1)th word; predicting said (j+1)th word based on said one or multiple partial analysis trees; obtaining a sentence

for a phrase including said predicted (j+1)th word and the probability value for said

sentence structure; and performing the above steps up to the last word of said sentence,

and selecting, as speech recognition results, a sentence structure and a word sequence

having maximum probability values.

5   For a better understanding of the present invention, together with other and further

features and advantages thereof, reference is made to the following description, taken in

conjunction with the accompanying drawings, and the scope of the invention that will be

pointed out in the appended claims.

Brief Description of the Drawings

10   Fig. 1 is a block diagram for the embodiment.

Fig. 2 is a diagram showing the configuration of a computer system according

to the embodiment.

Fig. 3 is a diagram for explaining word prediction according to the embodiment.

Fig. 4 is a diagram for explaining an example of word prediction according to

15   the embodiment.

Fig. 5 is a flowchart for explaining speech recognition according to the embodiment.

Fig. 6 is a diagram showing another example for explaining word prediction according to the embodiment.

5   Fig. 7 is a diagram showing an additional example for explaining word prediction according to the embodiment.

Fig. 8 is a diagram showing an example for explaining word prediction using a tri-gram model.

Fig. 9 is a diagram showing another example for explaining word prediction
10   using a tri-gram model.

Fig. 10 is a diagram showing an example for explaining word prediction using the method proposed by Chelba and Jelinek.

Fig. 11 is a diagram showing another example for explaining word prediction using the method proposed by Chelba and Jelinek.

Detailed Description of the Preferred Embodiments

The preferred embodiment of the present invention will now be described. Fig. 1 is a block diagram illustrating the configuration including a speech recognition apparatus, according to the embodiment. A word sequence W, generated as a sentence (a true sentence) by a block 101, is released as S (block 102). The released S is transmitted to an acoustic processor 111 constituting speech recognition means 110. The acoustic processor 111 converts the input S into a signal X, which it stores. The signal X is changed into a recognition result W' by a language decoder 112, which includes an acoustic model 113 that has learned the feature of a sound and a dictionary 114 in which text data prepared in advance through learning is stored. The sentence for the result W' is subsequently displayed (block 120).

Fig. 2 is a diagram showing an example system the employment of the speech recognition apparatus according to the embodiment. This system comprises a microphone 210, a computer 220 and a display device 230. The computer 220 includes a sound card 221, a memory 222 and a CPU 223. In the system in Fig. 2, a speaker's speech is received as an analog signal by the microphone 210. Thereafter, the sound card 221 of the computer 220 converts the analog signal into a digital signal that is stored in the memory 222. The acoustic model 113 and the dictionary 114, including

the language model, are also stored in the memory 222. Subsequently, the CPU 223

decodes the language based on the digital signal and the dictionary 114 stored in the

memory 222, and also interprets and executes a program for implementing a word

prediction method that will be described later. The obtained language is the

5   recognition result, and is displayed on the display device 230. This program is stored

in the memory 222.

In this system, the microphone 210 is a member separate from the computer

220, but it may be integrally formed with the computer 220 or the display device 230.

In other words, so long as a microphone for converting speech into equivalent electrical

10   energies is provided, any form can be employed. Furthermore, the recognition result is

displayed on the display device 230, e.g., a CRT; however, but the result can also be

transferred to and printed by a printer, or it can be stored on a flexible disk or another

storage medium.

In this embodiment, as an assumption for the word prediction method that will

15   be described below, proposed words are selected as the result of calculations that use

the acoustic model 113 for the speech signal obtained by the acoustic processor 111.

The following word prediction process is performed for these selected words, and the

speech recognition results are finally obtained.

The word prediction method using for this embodiment will now be described. As is described above, according to the present invention it is proposed that a partial analysis tree that has a modification relationship with a word to be predicted is predicted, and then, the partial analysis tree, i.e., a preceding word and/or word

5   sequence it is estimated is related to the next word, is employed to predict the next word. In other words, the next word is predicted by using the partial analysis tree that has a modification relationship with the word to be predicted.

This embodiment will be explained based on the example phrase "sara-ni sho-senkyoku no", which was used for explaining the tri-gram method and the method

10   proposed by Chelba and Jelinek. The phrase "sara-ni sho-senkyoku no" comprises the five words "sara-ni", "sho", "senkyo", "ku" and "no". Assuming "j" is used to represent the position of a word measured from the beginning of the phrase, "no" is the fifth word. Further, as shown in Figs. 3A to 3C, it is assumed that there are three structure types for the phrase that includes the fifth word "no". The sentence structure in this

15   case represents the modification relationship among the words. The three structures will now be described.

In Fig. 3A, while "sara-ni" does not modify "no", "no" is modified by "ku". This state is shown by using arrows; the arrow from "sara-ni" points to the word

following "no", while the arrow from "ku" points to "no". Since "sara-ni" forms a partial analysis tree and "sho-senkyoku no" forms another partial analysis tree, in the example in Fig. 3A the only partial analysis tree related to "no" is "sho-senkyoku no". It should be noted that the probability value for this structure is defined as 0.034.

5   In Fig. 3B, neither "sara-ni" nor "ku" modify "no". Therefore, the arrows from "sara-ni" and "ku" point to words following "no". The probability value for this sentence structure is defined as 0.001.

In Fig. 3C, instead of the "no" in Figs. 3A and 3B, the use of "wo", which has a similar sound, is predicted. The prediction of "wo" is instructed by the acoustic model 10   113. As for the sentence structure, as in Fig. 3A, while "sara-ni" does not modify "no", "no" is modified by "ku". The probability value for the case in Fig. 3C is defined as 0.028. And since the probability value in Fig. 3A is the highest, at this time the case represented by Fig. 3A, which has the maximum probability value, can be the proposed result for speech recognition.

15   The cases in Figs. 3A to 3C are merely examples used for explaining the embodiment. For example, when the fifth word is "wo", the same case as in Fig. 3B may be present, or a case where the fifth word is "to" instead of "no" or "wo" may be present. In any case, in Figs. 3A to 3C, the structure, including the j-th (fifth) word,

and the probability value are shown. It should be noted that the statement s[5][0] in Fig. 3 indicates that the fifth word is a target to be processed, and [0] means one of the words having a modification relationship is a target for the process.

Then, the sixth word is predicted. For this prediction, first, the sentence structure, including the sixth word, is specified. For the example in Fig. 3A, there are three available cases: a case where only "no" modifies the sixth word; a case where both "sara-ni" and "no" modify the sixth word; and a case where "sara-ni" and "no" do not modify the sixth word. The sixth word is predicted for the respective three cases. These three cases are shown in Figs. 3(a-1) through (a-3). In this embodiment, before the sixth word is predicted, the sentence structure, including the sixth word, is specified.

In the dictionary 114, the appearance frequency of a predetermined word relative to another predetermined word and/or word sequence is written based on text data that has been learned. For example, assuming that sentence "sho-senkyoku no" has appeared in the text data n times and has been followed by "donyu" m times, the frequency appearance for "donyu" relative to "sho-senkyoku no" is m/n. When two partial analysis trees of "sara-ni" and "sho-senkyoku no" are employed to predict "donyu", the frequency whereat "donyu" appears after "sara-ni" and "sho-senkyoku no"

must be taken into account. That is, assuming that, in the text data, a sentence

including "sara-ni" and "sho-senkyoku no" appeared n' times and thereafter the word

"donyu" appeared m' times, the appearance probability for "donyu" relative to "sara-ni"

and "sho-senkyoku no" is m'/n'. At this time, according to the empirical rule, very

5   frequently "sara-ni" will modify a declinable word, such as a verb or an adjective, and

will seldom modify an indeclinable word, such as a noun. Thus, since the appearance

frequency m' of noun "donyu" is very small, the probability value when "donyu" is

predicted by using two partial analysis trees "sara-ni" and "sho-senkyoku no" is

considerably smaller than the probability value obtained when "donyu" is predicted

10   merely by using "sho-senkyoku no". In other words, it is not preferable for "sara-ni" to

be taken into account for the prediction of "donyu".

Therefore, when "no" is used to predict "donyu", the probability value for the

phrase "sara-ni sho-senkyoku no donyu" is greater than the probability value for this

sentence when "sara-ni" and "no" are employed to predict "donyu".

15   In this embodiment, Figs. 3(a-1) and (a-2) have been explained, and the

probability value is calculated in the same manner for the case in Fig. 3(a-3). Further,

the prediction process is performed in the same manner, up to the last word of the

sentence.

The word prediction processing for the case in Fig. 3(a-1) will now be described while referring to Figs. 4A to 4C. In Fig. 4A, the state in Fig. 3(a-1) is shown. In this state, a partial analysis tree having a modification relationship with the next word "?" (the sixth word in this case) is specified. In this case, the partial analysis tree "sho-

5   senkyoku no" modifies the sixth word, while the sixth word is not modified by the partial analysis tree "sara-ni". This modification is shown in Fig. 4B. That is, the arrow from "sara-ni" points to a word following the sixth word, and indicates that no modification has been established between the sixth word and "sara-ni". The arrow from "no" in "sho-senkyoku no" points to the sixth word "?", and indicates that the

10   word sequence "sho-senkyoku no" modifies the sixth word.

As is described above, after the sentence structure, including the sixth word, has been predicted, "donyu" is predicted using the partial analysis tree "sho-senkyoku no", which has a modification relationship with the sixth word. Further, after the prediction of "donyu", as is shown in Fig. 4C, the sentence structure, to include "donyu", is

15   predicted. In other words, according to the case in Fig. 3(a-1), since "sara-ni", which probably reduces the prediction accuracy, is not taken into account, a high probability value can be obtained.

The word prediction method for this embodiment has been explained. Next, the processing for finally outputting the speech recognition results will be explained while referring to the flowchart in Fig. 5. According to this processing, as previously described, proposed words are selected as the results of calculations using the acoustic model 113 for speech signal acquired by the acoustic processor 111, and the narrowing of the selected words is further performed by the prediction.

In Fig. 5, which word is to be processed (S100) and which structure is to be processed (S101) are determined. The position of a word to be processed is represented by using "j", and a structure to be processed is represented by "i". Since the prediction is performed starting at the beginning of the sentence, the initial values of j and i are 0. The specific form of j and i can be easily understood by referring to Fig. 3.

Then, the structure of a sentence, including a word to be predicted, and its probability value are obtained (S102). In Fig. 5, s[j][] at S104 represents the sentence structure that includes the j-th word and the probability value. In the example in Fig. 3, first, s[5][0], i.e., the first sentence structure of the three, and its probability value are obtained for the fifth word. Since this sentence structure and the probability value are employed for the predication of the next word, these are enumerated relative to s[j+1][] (S102). In the example in Fig. 3, first, Fig. 3(a-1) is enumerated for s[6][].

When there are multiple sentence structures, the process at s102 is performed

for all of them. To do this, the process at S103, where i=i+1, and the process at S104,

for determining whether all s[j][] are examined, are performed.

When the process at S102 has been completed for all the structures, the same

5   process is performed for the next word, which is defined as j=j+1 (S105). When j=j+1

is not the last word of the sentence, the process sequence from S101 is performed.

When j=j+1 is the last word, the sentence structure and the word sequence having the

maximum probability value are selected from s[j][], and are displayed on the display

device 230. This sentence structure can be displayed by using arrows to indicate

10   modifications, or as a partial analysis tree structure.

In the above embodiment, the present invention is carried out on a personal

computer. However, the present invention can be provided as a storage medium in

which a predetermined program is stored, or a transmission apparatus for transmitting a

program.

15   The present invention will now be described based on an example phrase "nani-

ga ima seiji-no saisei-no tame-ni". The phrase "nani-ga ima seiji-no saisei-no tame-ni"

consists of nine words, "nani", "ga", "ima", "seiji", "no", "saisei", "no", "tame" and "ni",

and is constituted by three partial analysis trees "nani-ga", "ima" and "seiji-no saisei-no tame-ni".

In the state in Fig. 6A, the word prediction up to "nani-ga ima seiji-no saisei-no tame-ni" is completed. As is described above, this phrase is formed of three partial analysis trees, "nani-ga", "ima" and "seiji-no saisei-no tame-ni". As for the partial analysis tree "nani-ga", it has been predicted that "nani" modifies "ga". In other words, the word modified by the partial analysis tree "nani-ga" is unknown. This state is understood because the arrow from "ga" in Fig. 6A points to "?". Further, the words modified by the partial analysis trees "ima" and "seiji-no saisei-no tame-ni" are also unknown.

Based on the state in Fig. 6A, the partial analysis tree that modifies the next word (the tenth word in this example) is predicted. In this example phrase, it is predicted or specified that all of the three partial analysis trees, "nani-ga", "ima" and "seiji-no saisei-no tame-ni", modify the tenth word. This modification is shown in Fig. 6B. That is, the arrows from "ga" in "nani-ga", "ima", and "ni" in "seiji-no saisei-no tame-ni" point to the tenth word.

As is described above, when the sentence structure, to include the tenth word, has been specified, the tenth word is predicted. That is, since all three partial analysis

trees ("nani-ga", "ima" and "seiji-no saisei-no tame-ni") modify the word to be predicted, all of these are considered to predict "hitsuyo".

According to the method proposed by Chelba and Jelinek, "hitsuyo" is predicted using "tame" and "ni", while in this embodiment "nani-ga", which is useful information for predicting "hitsuyo", is also employed, the prediction accuracy in this embodiment is higher.

Up to now, Japanese phrases have been employed as examples. An explanation will now be given using an English phrase. One of differences between Japanese and English is that the direction of the modification in Japanese is constant, whereas it is not in English. When this embodiment is used for a language, such as English, where the direction of modification is not constant, only a partial analysis tree having a modification relationship with the next word and the direction of the modification need be specified, and the partial analysis tree having the modification relationship need only be employed to predict the next word.

Assume as an example that "after" is predicted from "the contact ended with a loss". The phrase "the contact ended with a loss" consists of six words, "the", "contact", "ended", "with", "a" and "loss". Further, "the contact" forms one partial analysis tree, and "ended with a loss" forms another partial analysis tree.

Fig. 7A is a diagram showing the state wherein the prediction of words up to "the contact ended with a loss" is completed. As is described above, this phrase consists of two partial analysis trees "the contact" and "ended with a loss". As indicated by arrows in Fig. 7A, "the" in the partial analysis tree "the contact" modifies

5   "contact". In the partial analysis tree "ended with a loss", "a" modifies "loss", "loss" modifies "with" and "with" modifies "ended". As is described above, the modification in English has two directions: from front to rear and from rear to front.

Based on the state in Fig. 7A, the partial analysis tree related to the next word "?" (the seventh word in this case) is predicted. In other words, it is predicted that, as

10   is shown in Fig. 7B, the seventh word modifies "ended". Since "ended" is included in the partial analysis tree "ended with a loss", the seventh word is predicted based on the modification relationship with "ended with a loss". Then, as is shown in Fig. 7C, "after" is predicted from the partial analysis tree "ended with a loss".

Example Experiment

15   A model consisting of approximately 1000 sentences was prepared based on a newspaper article. An experiment for obtaining an entropy was conducted for this model using the method of this embodiment. The following results were obtained.

This Embodiment: 4.05 [bit]

tri-gram: 4.27 [bit]

The value of 4.05 [bit] in this embodiment corresponds to a selection for which 16.6 words were used, and the value of 4.27 [bit] corresponds to a selection for which 19.3 words were used. Therefore, it was confirmed that the word prediction accuracy was improved when this embodiment was used.

As is described above, according to the present invention, the sentence structure, to include a word to be predicted, is specified, and the prediction of the word is performed using a word or a word sequence having a modification relationship with the word to be predicted. Since the modification relationship is useful information for the word prediction, the word prediction accuracy is increased.

If not otherwise stated herein, it is to be assumed that all patents, patent applications, patent publications and other publications (including web-based publications) mentioned and cited herein are hereby fully incorporated by reference herein as if set forth in their entirety herein.

Although illustrative embodiments of the present invention have been described herein with reference to the accompanying drawings, it is to be understood that the

invention is not limited to those precise embodiments, and that various other changes and

modifications may be affected therein by one skilled in the art without departing from the

scope or spirit of the invention